atmosphere

MDPI

Article

Contribution of Atmospheric Factors in Predicting Sea Surface Temperature in the East China Sea Using the Random Forest and SA-ConvLSTM Model

Qiyan Ji [1], Xiaoyan Jia [1,*], Lifang Jiang [2], Minghong Xie [1], Ziyin Meng [1], Yuting Wang [1] and Xiayan Lin [1]

[1] Marine Science and Technology College, Zhejiang Ocean University, Zhoushan 316022, China; jiqiyan@zjou.edu.cn (Q.J.); xieminghong@zjou.edu.cn (M.X.); mengziyin@zjou.edu.cn (Z.M.); wangyuting@zjou.edu.cn (Y.W.); linxiayan@zjou.edu.cn (X.L.)

[2] South China Sea Marine Forecast and Hazard Mitigation Center, Ministry of Natural Resources, Guangzhou 510310, China; jlf@hyyb.org

* Correspondence: s20070700001@zjou.edu.cn

Abstract: Atmospheric forcings are significant physical factors that influence the variation of sea surface temperature (SST) and are often used as essential input variables for ocean numerical models. However, their contribution to the prediction of SST based on machine-learning methods still needs to be tested. This study presents a prediction model for SST in the East China Sea (ECS) using two machine-learning methods: Random Forest and SA-ConvLSTM algorithms. According to the Random Forest feature importance scores and correlation coefficients R, 2 m air temperature and longwave radiation were selected as the two most important key atmospheric factors that can affect the SST prediction performance of machine-learning methods. Four datasets were constructed as input to SA-ConvLSTM: SST-only, SST-T2m, SST-LWR, and SST-T2m-LWR. Using the SST-T2m and SST-LWR, the prediction skill of the model can be improved by about 9.9% and 9.43% for the RMSE and by about 8.97% and 8.21% for the MAE, respectively. Using the SST-T2m-LWR dataset, the model's prediction skill can be improved by 10.75% for RMSE and 9.06% for MAE. The SA-ConvLSTM can represent the SST in ECS well, but with the highest RMSE and AE in summer. The findings of the presented study requires much more exploration in future studies.

Keywords: Random Forest; SA-ConvLSTM; East China Sea; sea surface temperature; prediction

check for updates

Citation: Ji, Q.; Jia, X.; Jiang, L.; Xie, M.; Meng, Z.; Wang, Y.; Lin, X. Contribution of Atmospheric Factors in Predicting Sea Surface Temperature in the East China Sea Using the Random Forest and SA-ConvLSTM Model. *Atmosphere* **2024**, *15*, 670. https://doi.org/10.3390/atmos15060670

Academic Editor: Stephan Havemann

Received: 28 April 2024
Revised: 23 May 2024
Accepted: 28 May 2024
Published: 31 May 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The ocean covers about 75% of the Earth's surface, and its thermal energy plays a crucial role in regulating the global climate and ecosystems. Sea surface temperature (SST) is one of the most important ocean parameters with significant impacts on global climate and marine ecosystems [1–4]. Accurate and effective forecasting of SST is important for marine weather services, climate prediction, marine fisheries production, and marine environmental protection [5–7]. The East China Sea (ECS) is a shelf sea in the northwestern Pacific Ocean, located to the west of mainland China, south of the Yellow Sea, and north of the South China Sea [8]. The SST of the ECS is influenced by many factors such as solar radiation, monsoon, East China Sea shelf circulation, Kuroshio intrusion, runoff, tides, etc. [9]. These factors are highly stochastic and uncertain, and pose many challenges to predicting the SST accurately. Researchers and scientists have proposed many methods for predicting SST, which can be divided into two categories: one is based on numerical ocean models, and the other is based on data-driven methods.

The numerical ocean model is a powerful tool for marine science research and operational oceanography, and it is commonly used to simulate the physical, chemical, and biological processes of the ocean [10]. Modern numerical ocean models are mainly based on complex physical equations, which are difficult to solve without certain approximations

and simplifications. SST prediction by numerical ocean models usually requires precise initial and boundary conditions and atmospheric forcing [11]. Many typical numerical ocean models, such as HYCOM, ROMS, POM etc., have been widely used to simulate and predict not only the SST but also the three-dimensional temperature field of the ocean [12]. Gao et al. used HYCOM to simulate the SST of the Tropical and North Pacific based on the different combinations of two air–sea flux data sets (COADS and ECMWF) and two bulk parameter formulas (non-constant and constant) [13]. Zhang et al. used the ROMS to simulate the three-dimensional sea temperature in the northern Yellow Sea and analyzed its distribution characteristics and variations [14]. Based on the POM, Zhao et al. constructed a three-dimensional baroclinic assimilation model to model to simulate the current and temperature field in the Bohai Sea, Yellow Sea, and East China Sea [15]. However, strong nonlinear instability, huge computational expense, low reusability efficiency, and high coupling costs have gradually become the main bottlenecks for the further development of numerical ocean modeling. The parameterization schemes of physical and dynamical processes in numerical ocean models, inaccurate initial and boundary conditions, and atmospheric forcing can easily lead to numerical model prediction errors and affect simulation and prediction accuracy.

In recent years, data-driven methods have also been widely used to build SST prediction models by learning the data variability rule from historical SST observations [16]. Data-driven methods rely primarily on observational data rather than solving complex physical governing equations. They are simpler, less computationally intensive, and more likely to achieve higher predictive performance than numerical models. Data-driven methods include traditional statistical models, shallow neural network models, and deep learning models [17–20]. Traditional statistical models such as Markov and regression models are first employed to predict SST. Xue et al. constructed a seasonally varying Markov model to predict SST and sea level in the tropical Pacific [21]. Laepple et al. used a regression model to predict SST and found a correlation between SST and the number of hurricanes [22]. Kug et al. developed a regression model for the dynamic prediction of monthly SST in the Indian Ocean based on the lagged relationship between SST and Nino 3 [23]. The above-mentioned statistical models can predict the evolution of the SST based on specific observational data, but it is difficult to further improve the accuracy of the forecasts due to data and methodological limitations [24]

As increasing SST data have become available, researchers have been trying to use new data-driven methods to predict SST. Shallow neural network models have become increasingly popular due to their high degree of flexibility in fitting SST data [25,26] and have shown substantial improvements in prediction accuracy compared to traditional statistical models. Tangang et al. used a neural network to predict SST anomalies in the equatorial Pacific Niño region, achieving a correlation skill of 0.8, demonstrating the feasibility of neural networks to capture nonlinear relationships [27]. Wu et al. developed a nonlinear prediction system for tropical Pacific SST anomalies using a multilayer neural network approach, which improved the correlation coefficient by 0.1–0.14 compared to a linear regression model [28]. Tripathi et al. used an artificial neural network to predict SST anomalies in a specific region of the Indian Ocean with high accuracy [29]. Aparna et al. (2018) proposed a three-layer neural network to predict SST at a given location with a prediction error of ± 0.5 °C [30]. Gupta et al. evaluated the prediction ability of different methods based on specific training, regression, and artificial neural networks and found that the neural networks showed a better RMSE value of 1.3 °C than the other methods [31]. However, many shallow neural networks cannot fully utilize the large amount of historical SST data to train prediction models due to their limited learning capacity and relatively simple structure [32].

In recent years, deep learning models have gained popularity in predicting SST due to their powerful learning and modeling capabilities [9]. Zhang et al. used a Long Short-Term Memory (LSTM) model to forecast SST in the Bohai Sea [33]. They inputted daily, weekly, and monthly SST data and achieved an RMSE value of less than 1.13 °C. Xiao et al.

developed an LSTM-AdaBoost model to predict the SST in the East China Sea region for the next 10 days, and the model achieved a high correlation coefficient of 95% [34]. He et al. proposed an SST prediction model based on the EMD-GRU method and found that the EMD-GRU model reduced mean square error (MSE) and mean absolute error (MAE) to different degrees compared with LSTM and GRU models [35]. Zheng et al. proposed a deep learning model for predicting SST in the equatorial eastern Pacific Ocean, which includes a convolutional layer, a max-pooling layer, and an up-sampling layer [20]. The experimental results show an RMSE value of less than 0.9 °C. Ham et al. proposed a model based on CNN to predict the Nino 3.4 index for the next 23 months and found that the CNN model could predict ENSO 1.5 years in advance with a correlation coefficient of up to 90% [36]. However, these methods only consider the temporal evolution of the SST data without taking into account its spatial correlation.

To address the above problem, Shi et al. transformed the precipitation proximity prediction into a space–time series prediction problem and proposed the ConvLSTM model, which was subsequently used in SST prediction [37]. Xiao et al. used the ConvLSTM model to predict the SST in the East China Sea for the next 10 days, and the results showed that the proposed model outperformed the SVR model and the LSTM mode [16]. Zhang et al. proposed the M-ConvLSTM model for three-dimensional temperature prediction, which showed strong prediction performance with a coefficient of determination R^2 greater than 0.84 at different water depths [38]. Although the ConvLSTM model is effective in predicting SST, it is limited in achieving long-range spatial dependence by stacking convolutional layers because the effective receptive field is much smaller than the theoretical receptive field [39]. To address the above issues, Lin et al. proposed a self-attentive memory module (SAM) for ConvLSTM, which is embedded in the ConvLSTM model to enable it to effectively capture long-range spatial dependencies, referred to as SA-ConvLSTM [40]. Therefore, this paper will use the SA-ConvLSTM model to predict the SST in the East China Sea.

Since most current data-driven approaches ignore physical mechanisms, both the methodology and data are crucial for any data-driven SST prediction model. Although there has been much research on innovative methodologies, further investigation is required to determine the impact of the dataset on prediction accuracy. In addition to the essential SST data, are there other datasets that have a physical relationship with SST that could affect SST forecast accuracy when using a data-driven method? Atmospheric forcings have a strong physical relationship with SST variability and are often considered an essential input for predicting SST using numerical oceanic models. It is worth considering how these data can be used to improve the accuracy of the SST forecasts using the data-driven method that has been selected, namely the SA-ConvLSTM model.

This paper aims to investigate the influence of atmospheric factors on the accuracy of SST predictions using SA-ConvLSTM. The study considers various atmospheric factors, including the 2 m air temperature, sea surface solar shortwave radiation flux, sea surface longwave radiation flux, sea surface latent heat flux, sea surface sensible heat flux, and the net heat flux, wind of 10 m. These factors are considered to be the most important atmospheric factors that can affect the sea surface temperature. The feature importance score of the Random Forest model was used to assess the contribution of these atmospheric factors and select the critical factors to construct different datasets. The SA-ConvLSTM model is used to predict the SST in the East China Sea, and the impact of different datasets on the prediction results of the SA-ConvLSTM model will be analyzed. The remainder of the paper is organized as follows. Section 2 focuses on the data sources and pre-processing. Section 3 provides a detailed account of the methodologies employed in this study and the construction of the model. Section 4 presents the experimental results. Section 5 offers a discussion of the findings. Section 6 presents the conclusion.

2. Data

The high-resolution satellite remotely sensed SST data used in this paper is the Operational Sea Surface Temperature and Ice Analysis (OSTIA) [41]. OSTIA is a daily global SST product developed by the UK Met Office with a resolution of 1/20°. These data are derived from AATSR, SEVIRI, AVHRR, AMSR, and TMI measurements with a root mean square error of less than 0.6 °C. The data can be found on the following website: https://www.ncei.noaa.gov/data/oceans/ghrsst/L4/GLOB/UKMO/OSTIA/ (accessed on 11 November 2021). The paper uses data with a spatial range of 22° N~33° N, 120° E~131° E, and a temporal range of 2010 to 2020. The training data set consists of data from 2010 to 2019, while the test data set consists of data from 2020.

The atmospheric factors data used in this paper include 2 m air temperature (T2m), sea surface solar shortwave radiation flux (*SWR*), sea surface longwave radiation flux (*LWR*), sea surface latent heat flux (LHF), sea surface sensible heat flux (*SHF*), U-component of 10 m wind and V-component of 10 m wind (10 m-wind) and were provided by ERA5. ERA5 is the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of global climate data. ERA5 combines model data with observations to provide a numerical interpretation of the recent climate. It offers hourly data on atmospheric, terrestrial, and oceanic climate variables at a resolution of 1/4°. The data can be found here: https://cds.climate.copernicus.eu/ (accessed on 20 December 2021). The paper uses data with a spatial range of 22° N~33° N, 120° E~131° E, and a temporal range from 2010 to 2020. The training dataset consists of data from 2010 to 2019, while the test dataset consists of data from 2020. The sum of heat flux into or out of the water, which is referred to as the net heat flux $\mathcal{Q}net$, is also considered to be a factor that can have an impact on sea surface temperature. Since we only consider the influence of atmospheric factors, the heat flux carried by ocean currents to the regional sea is not considered in this paper. Refer to the heat budget equation in [42]; the $\mathcal{Q}net$ is calculated with the following equation:

$$\mathcal{Q}net = SWR + LWR + LHR + SHF \quad (1)$$

3. Method

3.1. Random Forest

Random Forest (RF) is a non-parametric supervised learning method for classification and regression. It was first proposed by Breiman and can be regarded as a special type of bagging algorithm [43,44]. Random Forest is widely used as a classifier in remote sensing. It can be successfully used to select and rank the variables with the greatest ability to discriminate between the target classes [45]. The Random Forest is also used in various other research projects, including the environmental impact assessment of multi-source solid waste [46], the prediction of perennial ryegrass biomass [47], the identification and control of land subsidence and uplift [48], and the predicting of SST [49,50]. As Random Forest can be used to select and rank the key variables, it is well suited to selecting and ranking the key atmospheric factors that can improve the accuracy of the SST prediction. Random Forest consists of multiple decision trees, each of which is trained on independent data, and a final prediction is obtained by voting or averaging. The algorithm flow is shown in Figure 1. In this paper, we use Random Forests to evaluate the importance of the features that can affect the SST, and then we select the most important features based on their score in the feature importance evaluation.

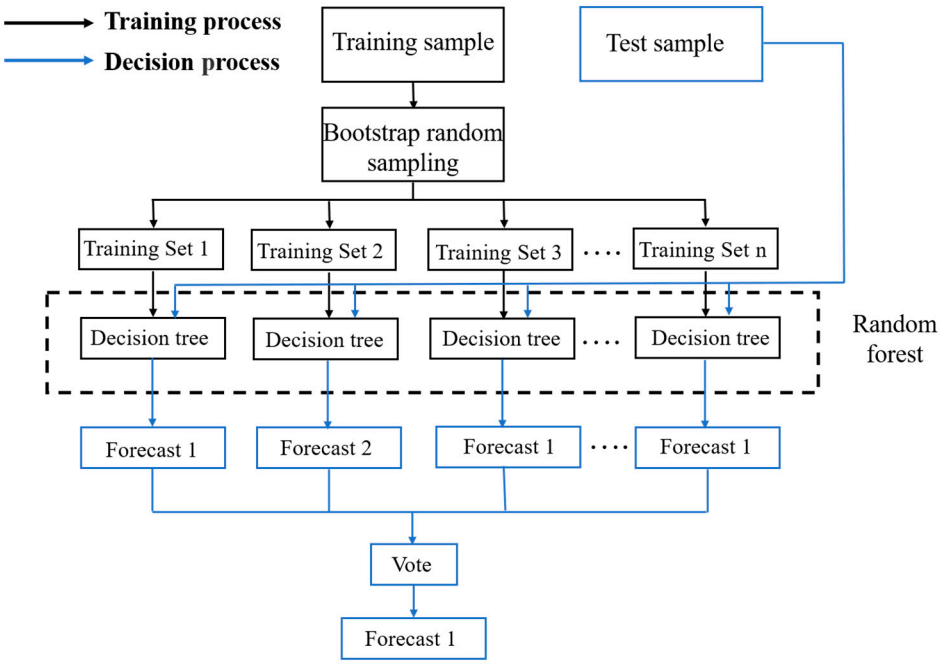


Figure 1. Flow chart of the Random Forest algorithm.

3.2. SA-ConvLSTM Model

The SA-ConvLSTM model was proposed by Lin et al., which can capture the long-range spatial dependence of the data well [40]. The SA-ConvLSTM model is a combination of the self-attention module and ConvLSTM, as shown in Equation (2). However, Lin et al. discovered that this basic model did not yield satisfactory experimental results, prompting them to make improvements to the self-attention module [40].

$$\begin{cases} X_t = SA\left(\overline{X}_t\right) \\ \hat{H}_{t-1} = SA\left(\hat{h}_{t-1}\right) \\ i_t = \sigma\left(W_i * \left[\hat{H}_{t-1}, X_t\right] + b_i\right) \\ f_t = \sigma\left(W_f * \left[\hat{H}_{t-1}, X_t\right] + b_f\right) \\ o_t = \sigma\left(W_o * \left[\hat{H}_{t-1}, X_t\right] + b_o\right) \\ g_t = tanh\left(W_c * \left[\hat{H}_{t-1}, X_t\right] + b_c\right) \\ C_t = f_t \odot C_{t-1} + i_t \odot g_t \\ \hat{H}_t = o_t \odot tanh(C_t) \end{cases} \quad (2)$$

where SA represents the self-attention module, X_t and $\hat{H}_{t-1}$ represent the features that are aggregated through the self-attention module. Figure 2 shows the structure of the SA-ConvLSTM model, in which the self-attention memory module (SAM) is embedded in the ConvLSTM model. If the SAM module is removed, the SA-ConvLSTM model becomes a standard ConvLSTM model. The internal calculation formulas for the self-attention memory module are introduced in Section 3.2.2.

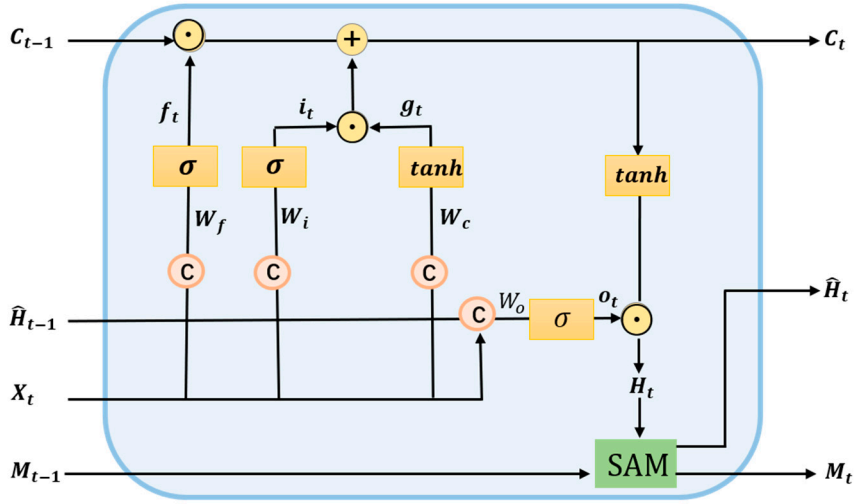


Figure 2. Structure of the SA-ConvLSTM model, where SAM is a self-attention memory module.

3.2.1. ConvLSTM Model

The ConvLSTM model consists of a memory unit and three memory gates. The memory unit can be considered to be a memory update gate. The structure of the memory unit is shown in Figure 3, and the equation is shown below:

$$\begin{cases} \bar{i}_t = \sigma\left(\overline{W}_i * \left[\hat{h}_{t-1}, \overline{X}_t\right] + \bar{b}_i\right) \\ \bar{f}_t = \sigma\left(\overline{W}_f * \left[\hat{h}_{t-1}, \overline{X}_t\right] + \bar{b}_f\right) \\ \bar{g}_t = tanh\left(\overline{W}_c * \left[\hat{h}_{t-1}, \overline{X}_t\right] + \bar{b}_c\right) \\ C_t = \bar{f}_t \odot C_{t-1} + \bar{i}_t \odot \bar{g}_t \\ \bar{o}_t = \sigma\left(\overline{W}_o * \left[\hat{h}_{t-1}, \overline{X}_t\right] + \bar{b}_o\right) \\ \hat{h}_t = \bar{o}_t \odot tanh(C_t) \end{cases} \quad (3)$$

where $\hat{h}_{t-1}$ denotes the hidden state at the previous moment and $\overline{X}_t$ denotes the current input value. σ and tanh are activation functions, * denotes the convolution operation, and $\odot$ denotes the Hadamard product. $\bar{f}_t, \bar{i}_t, \bar{g}_t, \bar{o}_t$ denote forget gate, input gate, update gate, and output gate. $\overline{W}_f, \overline{W}_i, \overline{W}_c, \overline{W}_o$ and $\bar{b}_f, \bar{b}_i, \bar{b}_c, \bar{b}_o$ represent the weight matrix and bias matrix, respectively. C_{t-1} and C_t represents the unit state at the previous time and the unit state at the current time.

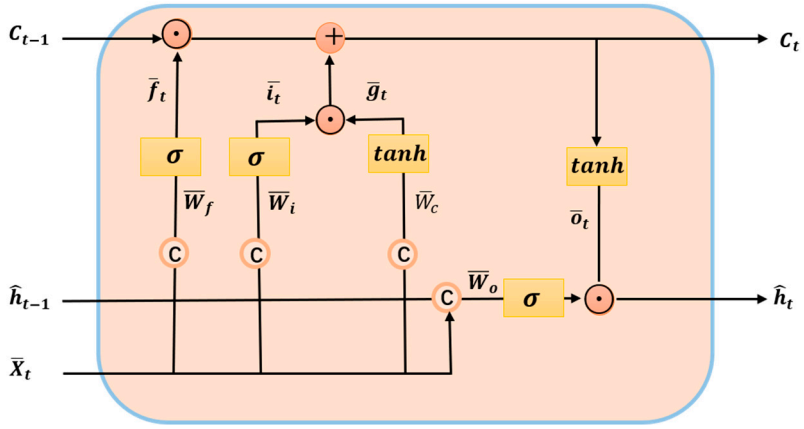


Figure 3. ConvLSTM memory unit structure.

3.2.2. Self-Attention Memory Module

The basic structure of the self-attention memory module (SAM) is presented in Figure 4, from which the SAM accepts two main inputs, namely the feature H_t of the current time step and the memory M_{t-1} of the previous step. It consists of three main parts: feature aggregation, memory updating, and output.

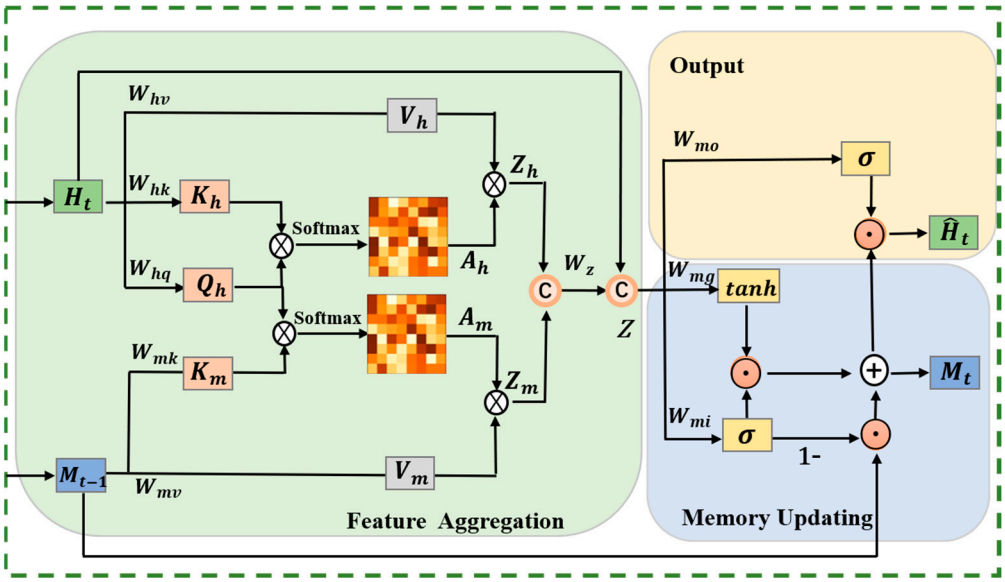


Figure 4. Structure of the Self-Attention Memory Module.

(1) Feature aggregation. The aggregated feature Z at each time step is formed by fusing Z_h and Z_m. Z_h is calculated by mapping the original feature H_t to different feature spaces as the query: $Q_h = W_{hq}H_t \in R^{\hat{C}\times N}$; the key: $K_h = W_{hk}H_t \in R^{\hat{C}\times N}$; and the value: $V_h = W_{hv}H_t \in R^{C\times N}$; where W_{hq}, W_{hk} and W_{hv} are the weight matrices of the 1×1 convolution kernel, C and $\hat{C}$ are the number of channels, and $N = H \times W$. Then, the similarity weight of the aggregated feature is calculated by the formula $A_h = Softmax(Q_h^T K_h)$ and finally the similarity scores of the aggregated feature are obtained by the formula $Z_h = V_h A_h$.The procedure for calculating Z_m is similar to that of Z_h, but the similarity score of the aggregated feature of Z_m is obtained by calculating the features of the current time step and the memory feature from the previous step. The final aggregated feature Z can be obtained with $Z = W_Z[Z_h; Z_m]$.

(2) Memory updating. The gating mechanism is used to update the memory M adaptively and enable SAM to capture long-range dependencies in the spatial and temporal domains. The aggregated feature Z and the original input H_t are used to generate the input gate i'_t and the fused feature g'_t. In addition, the forgetting gate is replaced with $1 - i'_t$ to reduce the parameters. The update process is shown below.

$$\begin{cases} i'_t = \sigma(W_{m;zi} * Z + W_{m;hi} * H_t + b_{m;i}) \\ g'_t = tanh\left(W_{m;zg} * Z + W_{m;hg} * H_t + b_{m;g}\right) \\ M_t = (1 - i'_t) \odot M_{t-1} + i'_t \odot g'_t \end{cases} \tag{4}$$

where $W_{m;zi}$, $W_{m;hi}$, $W_{m;hg}$ denote the weight matrices. $b_{m;i}$ and $b_{m;g}$ denote the bias matrices.

Compared with the original memory cell C in the ConvLSTM, which is updated only by convolutional operations, the proposed memory M is updated not only by convolutional operations but also by aggregation of features Z_t, obtaining global spatial correlations in time. Therefore, we consider that M_{t-1} can contain global past spatial-temporal information.

(3) Output. The output feature $\hat{H}_t$ of the self-attention memory module is a dot product between the output gate ${o'}_t$ and the update memory M_t, with the following equation.

$$\begin{cases} {o'}_t = \sigma(W_{m;zo} * Z + W_{m;ho} * H_t + b_{m;o}) \\ \hat{H}_t = {o'}_t \odot M_t \end{cases} \quad (5)$$

where $W_{m;zo}$, $W_{m;ho}$ denote the weight matrix. $b_{m;o}$ denotes the bias matrix.

3.3. Model Construction

In this paper, the computational resources employed are Python 3.8. The Random Forest model was implemented via the Sklearn library, whereas the SA-ConvLSTM model was developed using the Tensorflow and Keras frameworks. Finally, the SST prediction model was constructed using both Keras and the Sklearn toolkit (Figure 5). The optimizing parameters of the Random Forest model have been set up in detail as follows: the number of decision trees is 1000, the maximum depth of the tree is 30, and the maximum number of features is 7. The SA-ConvLSTM model has three SA-ConvLSTM layers and one Conv3D layer with convolution kernels of (3, 3) and (3, 3, 3), respectively, and the optimization algorithm is Adam. In addition, to prevent overfitting, we have added dropout layers and set them to 0.4 and 0.3, respectively, and the maximum number of iterations set to 40.

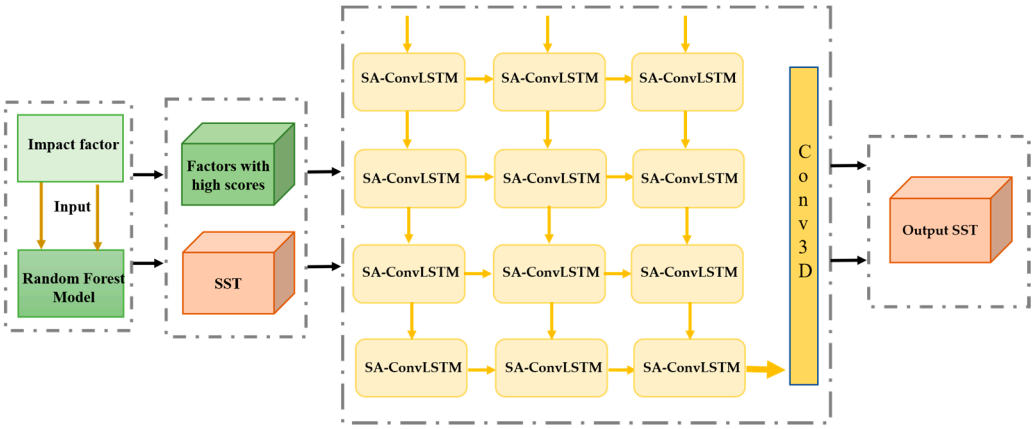


Figure 5. Structure of SST prediction based on Random Forest and SA-ConvLSTM model.

3.4. Evaluation Functions

To evaluate the SST prediction performance, this paper uses the root mean square error (*RMSE*), the mean absolute error (*MAE*), the absolute error (*AE*), the coefficient of determination (R^2), and the correlation coefficient (R) to compare and analyze the predicted SST with the actual SST data. The equations are as follows.

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(sst_o - sst_p\right)^2} \quad (6)$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}\left|sst_o - sst_p\right| \quad (7)$$

$$AE = \left|sst_o - sst_p\right| \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}\left(sst_o - sst_p\right)^2}{\sum_{i=1}^{m}\left(sst_o - \overline{sst_o}\right)^2} \quad (9)$$

$$R = \frac{\sum_{i=1}^{m}\left(var_i - \overline{var}\right)\left(sst_o - \overline{sst_o}\right)}{\sqrt{\sum_{i=1}^{m}\left(var_i - \overline{var}\right)^2}\sqrt{\sum_{i=1}^{m}\left(sst_o - \overline{sst_o}\right)^2}} \quad (10)$$

$$PBIAS = \frac{\sum_{i=1}^{m}(sst_p - sst_o)}{\sum_{i=1}^{m} sst_o} \times 100\% \quad (11)$$

where m is the total number of samples, sst_o and sst_p are the OSTIA data and the predicted value of the SST. The lower the value of RMSE, MAE, and AE, the more accurate the predicted result. The closer the R^2 is to 1, the better the fit between the predicted SST and the OSTIA data. The higher the R-value, the better the correlation. The Percent bias (PBIAS) measures the average tendency of the simulated data to be larger or smaller than observed counterparts. PBIAS can also help identify average model simulation bias (overprediction vs. underprediction) and incorporate measurement uncertainty.

To evaluate the improvement of the model using multi-element input dataset compared with a single-element input dataset, this paper uses the following formula:

$$Skill_{index} = \frac{\left|E_{index}^{ds} - E_{index}^{dm}\right|}{E_{index}^{ds}} \times 100\% \quad (12)$$

where E_{index}^{ds} represents the error index of the model's prediction using a single-element input dataset, E_{index}^{dm} represents the error index of the model's prediction using a multi-element input dataset, and the error index represents the RMSE value or MAE value.

4. Analysis of Results

4.1. Feature Importance Assessment for Random Forests

The SST in the ECS varies greatly from coastal to open sea, influenced by complex shelf circulations and monsoons. The mechanisms of SST variability are quite different in different regions, such as estuaries, coastal upwelling, and the Kuroshio Current. The mechanisms of SST variability are quite different in different regions, such as estuaries, coastal upwelling, and the Kuroshio Current. However, they are all influenced by atmospheric forcings such as the 2 m air temperature (T2m), sea surface solar shortwave radiation flux (*SWR*), sea surface longwave radiation flux (*LWR*), sea surface latent heat flux (LHF), sea surface sensible heat flux (*SHF*), U-component of 10 m wind and V-component of 10 m wind (10 m-wind). Since the data-driven methods could predict the future trend of SST based on a lot of SST data and ignore the physical mechanisms, not all these atmospheric forcings contribute equally to improving the accuracy of SST prediction. To select which kind of data could most improve the accuracy of SST, this study uses the Random Forest model to test the contribution of these data according to their feature importance assessment results in six typical locations L1~L6 (Figure 6). These sites are selected as follows: L1 at the Yangtze River estuary, where the SST has obvious seasonal variations. L2 and L3 are selected in the Kuroshio region, where temperature and salinity are influenced by the Kuroshio. L4 is selected in the northwest Pacific Ocean, which is far from the land, and the SST changes are relatively stable. L5 was chosen to be close to the Taiwan Strait, where temperature and salinity are influenced by the Taiwan warm current and the Min-Zhe coastal current [51,52]. L6 was chosen to be close to the Tsushima Strait, where temperature and salinity are influenced by the Tsushima Warm Current, a branch of the Kuroshio in the East China Sea region.

Figure 7 shows the prediction results of the Random Forest in 2020. In comparison with the OSTIA data, the Random Forest can accurately reproduce the long-term trend of the SST. The prediction results of Random Forest are not very accurate for the extremum, while all R^2 at six locations are above 94%, showing that the model built by Random Forest is reliable and stable. Both the R and PBIAS also show that the predicted SST result from Random Forest is well in agreement with the SST of OSTIA. The Random Forest feature importance score can then be used to select the key atmospheric forcings that affect SST variability. To enhance the accuracy of variable selection for predicting SST results, we compared both the Random Forest feature importance score and the correlation coefficient R (Figure 8). The results of the Random Forest forecast are more accurate when 2 m air

temperature and sea surface longwave radiation are taken into account, with 2 m air temperature being the most significant factor (Figure 8a). Conversely, 10 m wind speed, sea surface solar shortwave radiation, and heat flux have a lesser impact on the accuracy of the results. The sea surface longwave radiation and the 2 m air temperature have a strong positive correlation with the SST, with R values around 0.8 (Figure 8b). On the other hand, the R-value of the 10 m wind speed is mostly around −0.2, indicating a weak negative correlation with the SST. Although the histograms of the characteristic importance scores of the Random Forest and the histograms of R values differ, it can be observed that they have higher scores for longwave radiation and air 2 m temperature and lower scores for 10 m wind speed. In Section 4.2 of this paper, we selected longwave radiation (LWR) and 2 m air temperature (T2m) as influencing factors on SA-ConvLSTM forecast results and constructed four different datasets based on OSTIA SST, LWR, and 2 m air temperature data.

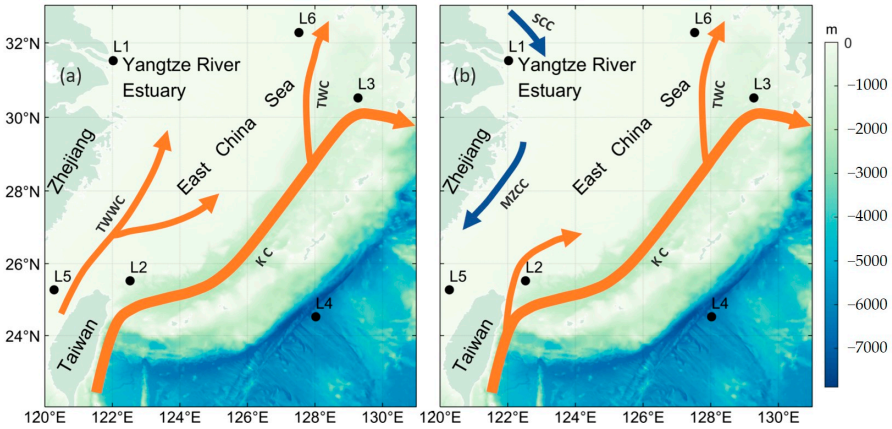


Figure 6. Schematic diagram of the East China Sea study area, where the black points are the six selected representative locations, where (**a**) shows the distribution of currents in summer and (**b**) the distribution of currents in winter [51,52]. SCC, MZCC, TWWC, TWC, and KC represent the Subei Coastal Current, the Min-Zhe (Fujian–Zhejiang) Coastal Current, the Taiwan Warm Current, the Tsushima Warm Current, and the Kuroshio. The blue line is the cold current, and the orange line is the warm current.

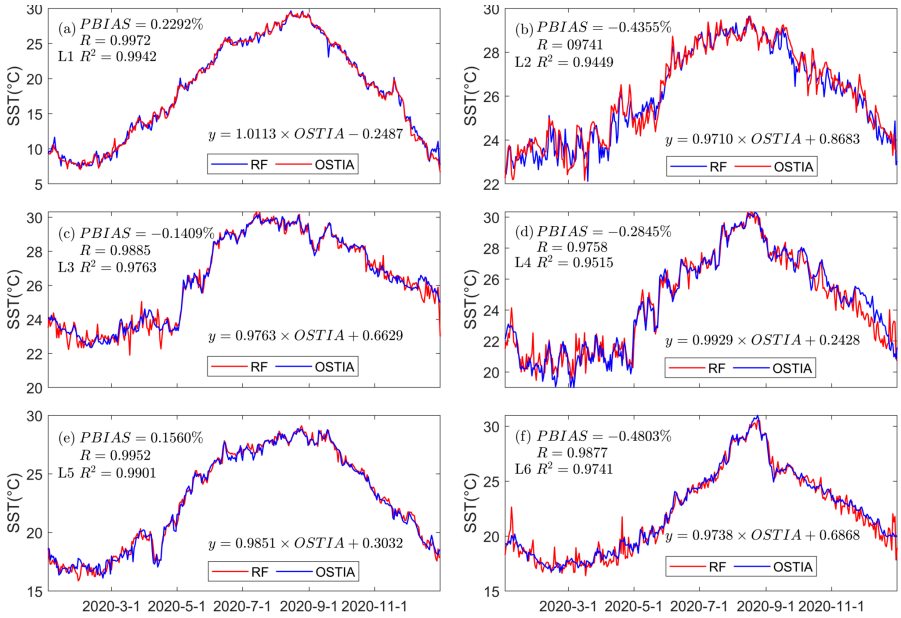


Figure 7. OSTIA data and prediction results of Random Forest (RF) for different locations, where (**a**–**f**) represent the sites of L1~L6, and the red line is the OSTIA data, the blue line is the prediction results of RF, and the horizontal coordinates represent the data.

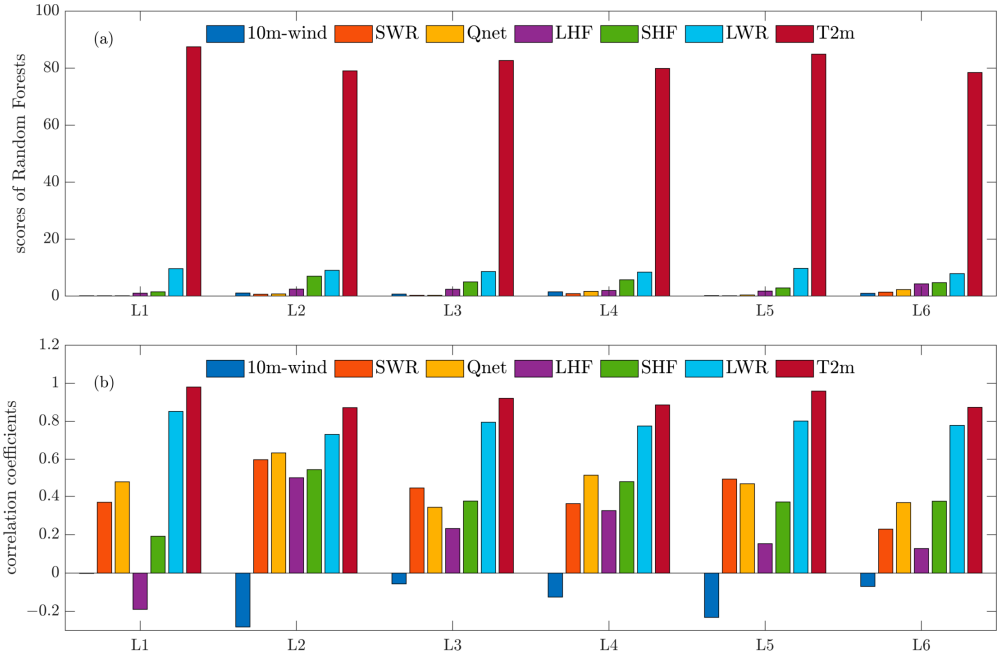


Figure 8. Characteristic importance assessment scores for Random Forests (**a**) and correlation coefficients (**b**) at six sites L1~L6. 10 m-wind is the wind speed of 10 m, SWR is the sea surface solar shortwave radiation, *Qnet* is the heat flux, LHF is the sea surface latent heat flux, SHF is the sea surface sensible heat flux, LWR is the sea surface longwave radiation and T2m is the 2 m air temperature.

4.2. Spatial Variation Analysis of Prediction Results from SA-ConvLSTM Models

Based on the results of the Random Forest selection and the correlation coefficient, four datasets were constructed as input to SA-ConvLSTM: SST-only, SST-T2m, SST-LWR, and SST-T2m-LWR. SST-only includes SST data only, SST-T2m includes both SST and 2 m air temperature data, SST-LWR includes SST and longwave radiation data, and SST-T2m-LWR includes SST, 2 m air temperature and longwave radiation data. To analyze the prediction accuracy of different input datasets, Figures 9–11 show the spatial variation of the seasonal mean predicted SST, AE, and RMSE. The four seasons are defined as follows: spring (March to May), summer (June to August), autumn (September to November), and winter (December to February). The predicted sea surface temperature (SST) of the four datasets has a spatial distribution similar to that of the OSITA data in all four seasons (Figure 9). The spatial characteristics indicate a general trend of low values in the northwest coastal seas and high values in the southeast open sea. The SA-ConvLSTM can represent the seasonal variation of SST in the ECS. The spatial distribution in spring and winter is consistent with OSTIA, while there are some differences in summer and autumn. Compared with OSTIA, the predicted SST by all four datasets in the Pacific Northwest region in summer and autumn is underestimated. The bias of the predicted result using the SST-only dataset is more apparent than the results of the other three datasets. The accuracy of the SST prediction can be improved by including T2m and LWR as additional input data.

The absolute error (AE) between predicted and observed values is low in spring and winter when using four different datasets as input, with most values around 0.2 °C (Figure 10). However, the AE between the predicted SST and OSTIA SST is largest in summer, regardless of which of the four datasets is used as input data, with an AE of about 1 °C in the Pacific Northwest region. The AE of all predicted SST in autumn is also evident in the Kuroshio region and coastal seas. The spatial distribution of AE differs significantly between the four seasons. During spring and winter, the AEs in coastal regions are higher than those in the open sea, indicating that the predicted temperature may be affected by coastal currents. During summer and autumn, the AE in the Kuroshio region is significantly

higher than that in the coastal region, suggesting that the predicted temperature may be influenced by the Kuroshio.

Figure 11 shows the spatial distribution of the root mean square error (RMSE) of the predicted SST using the four datasets. During summer, the RMSE is highest throughout the study area, with the largest RMSE observed in the region adjacent to Taiwan Island. The RMSE is lowest in winter, but the Yangtze River estuary, Zhejiang coastal seas, and the northeast sea of Taiwan have much higher RMSE than other parts of the study area. The spatial distributions of the RMSE in autumn are similar to those in winter, while the RMSE is clearly demarcated by the Kuroshio in spring, with high values in the northwestern area of the Kuroshio and low values in the southeastern area of the Kuroshio. In summer, the RMSE of datasets SST-T2m and SST-T2m-LWR are lower than those of SST-only and SST-LWR, which indicates that datasets including 2 m air temperature data perform better than those without. All four datasets have larger RMSE values and lower prediction accuracy in the summer in the Kuroshio regions, indicating that there is a significant correlation between the spatial distribution of model prediction accuracy and ocean circulations.

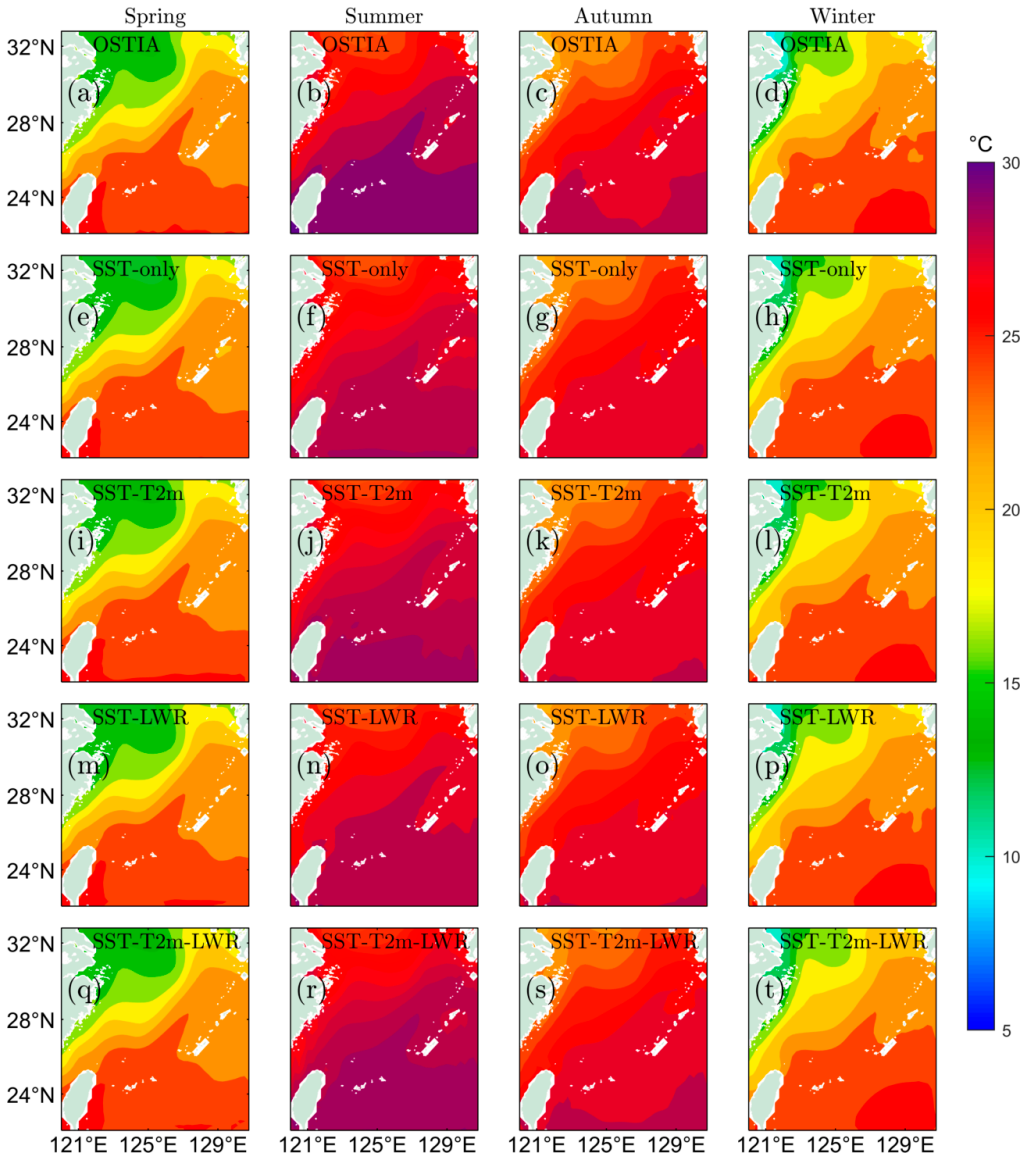


Figure 9. (**a**–**t**) Seasonal spatial distribution of OSTIA and the predicted SST from four constructed datasets: SST-only, SST-T2m, SST-LWR and SST-T2m-LWR.

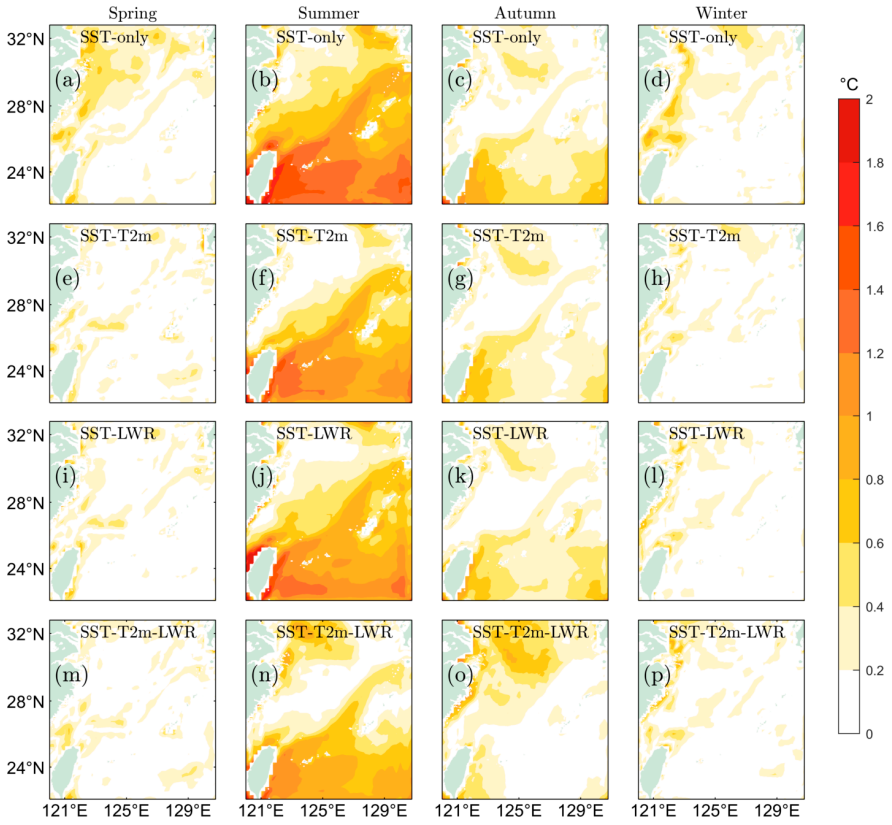


Figure 10. (**a**–**p**) Seasonal spatial distribution of AE of the predicted SST from four constructed datasets: SST-only, SST-T2m, SST-LWR and SST-T2m-LWR.

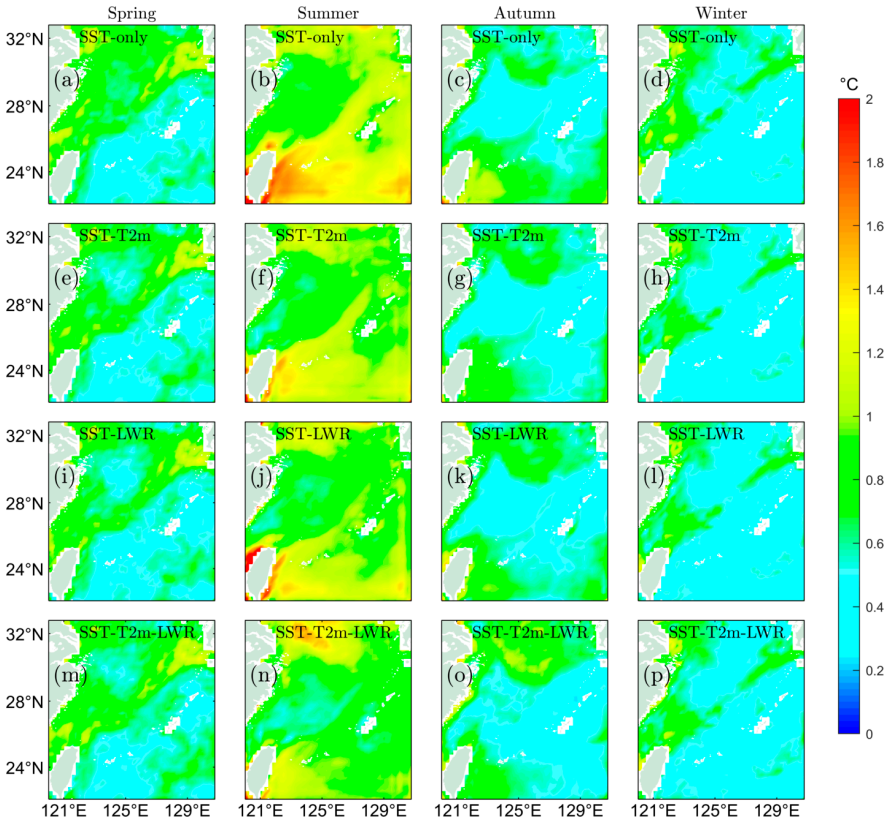


Figure 11. (**a**–**p**) Seasonal spatial distribution of RMSE of the predicted SST from four constructed datasets: SST-only, SST-T2m, SST-LWR and SST-T2m-LWR.

4.3. Temporal Variation Analysis of Prediction Results from SA-ConvLSTM Models

Table 1 presents the mean RMSE and MAE of the prediction results for four input datasets in 2020, along with the percentage improvement compared to the SST-only dataset. The RMSE and MAE of the prediction results using the SST-only dataset are the highest, with values of about 0.72 °C and 0.57 °C, respectively. Both the RMSE and MAE were reduced when using the SST-T2m, SST-LWR, and SST-T2m-LWR datasets. Compared to using only SST data as model training and input data for SA-ConvLSTM, adding T2m and LWR data separately can improve model performance. By adding the T2m and LWR data individually, the prediction skill of the model can be improved by about 9.9% and 9.43% for the RMSE and by about 8.97% and 8.21% for the MAE. However, when T2m and LWR data are added together, the prediction skill of the model using SST-T2m-LWR can be improved by 10.75% for RMSE and 9.06% for MAE. These results are slightly better than the results obtained by adding the T2m and LWR data separately. In descending order, both the RMSE and MAE of four datasets are SST-only, SST-LWR, SST-T2m, and SST-T2m-LWR. This shows that model prediction performance improves as the number of input datasets increases. Nevertheless, our experiments also show that when the number of input datasets increases from two to three types, there is no significant improvement in model performance. The Random Forest importance scores and R-value can be used as a reference for selecting additional input data to improve model prediction performance when using a data-driven method.

Table 1. The mean RMSE and MAE of four datasets in 2020 (unit: °C) and the percentage improvement.

	SST-Only	**SST-T2m**	**SST-LWR**	**SST-T2m-LWR**
MAE	0.5563	0.5064	0.5106	0.5059
RMSE	0.7221	0.6506	0.6540	0.6445
$Skill_{RMSE}$	0%	9.9%	9.43%	10.75%
$Skill_{MAE}$	0%	8.97%	8.21%	9.06%

To quantitatively assess the forecasting effectiveness of the model over the four seasons. Figure 12 displays the variation in the histogram of mean RMSE and MAE of prediction results for four input datasets across four seasons in 2020. In terms of seasonal variation, both the RMSE and MAE for the four datasets show similar seasonal variations, with the highest values in summer and the lowest values in winter. In summer, using the SST-only dataset resulted in the highest RMSE and MAE of about 1.01 °C and 0.88 °C, respectively, but using the SST-T2m-LWR dataset resulted in the lowest RMSE and MAE of about 0.82 °C and 0.70 °C, respectively. However, in winter, using the SST-only dataset resulted in the lowest RMSE and MAE of about 0.5 °C and 0.39 °C, respectively, but using the SST-T2m dataset resulted in the lowest RMSE and MAE of about 0.45 °C and 0.35 °C, respectively. In spring and autumn, the RMSE and MAE are about 0.6 °C and 0.5 °C, respectively. In terms of input datasets, the model's performance improvement due to T2m and LWR is most noticeable during the summer, with less significant improvements in other seasons. For the three seasons of spring, autumn, and winter, the RMSE differences among the four datasets are minimal. Similarly, the MAE differences among the four datasets are also small, except for summer.

Figure 13 shows the results of an analysis of the variation in the model's prediction errors using different input sets from month to month. The figure clearly shows that the maximum value of MAE and the minimum value of R^2 are obtained simultaneously. This relationship is particularly evident in summer. Furthermore, it was observed that the error curve in the prediction results increases as the month progresses, peaking in August when different input datasets are used. The model demonstrates a decrease in error as the month progresses after August. However, there is less improvement in prediction performance using the multi-input dataset compared to the single-input dataset before June and after October. In individual months, the model's predictions using two datasets are slightly

better than the model's predictions using three input datasets. For example, in December, the model's predictions using the SST-only input dataset have lower MAE values than the model's predictions using the SST-T2m-LWR input dataset. From June to October, the MAE value gradually decreases as the number of input datasets increases. This indicates that the model's prediction performance improves with more input datasets.

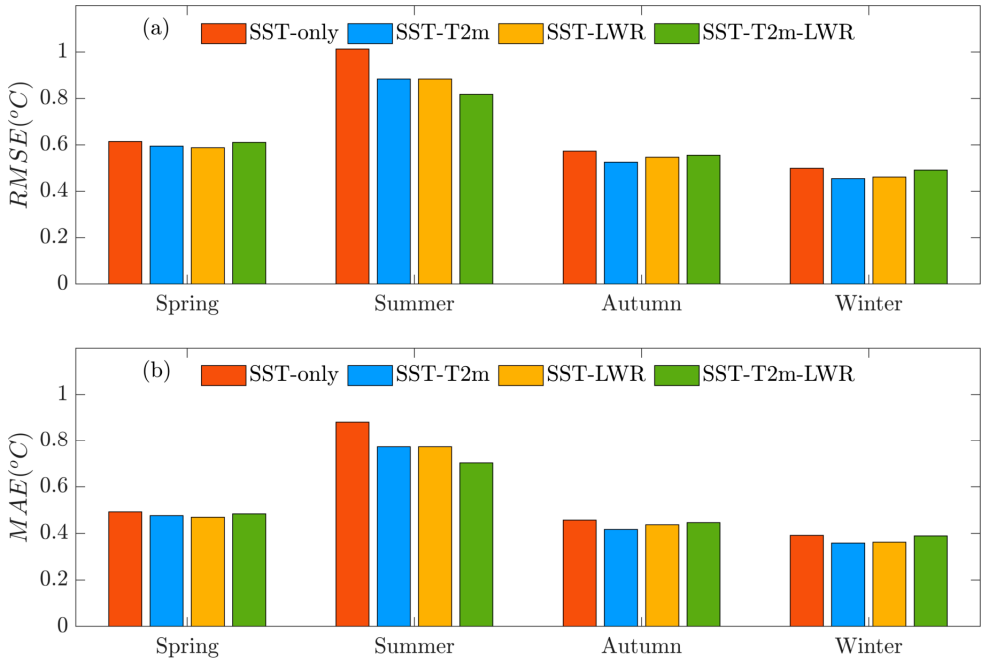


Figure 12. Seasonal variation of RMSE (**a**) and MAE (**b**) of prediction SST of four constructed datasets: SST-only, SST-T2m, SST-LWR and SST-T2m-LWR.

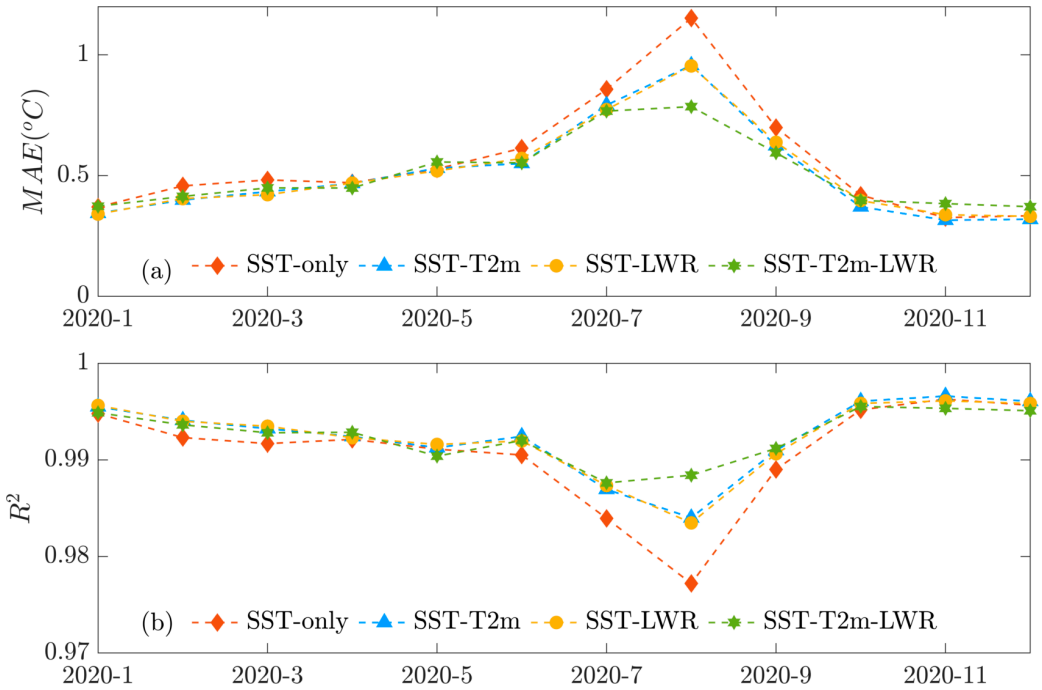


Figure 13. Seasonal monthly variation of MAE (**a**) and R^2 (**b**) of prediction SST of four constructed datasets: SST-only, SST-T2m, SST-LWR and SST-T2m-LWR.

5. Discussion

Numerical ocean modeling and data-driven methods are the two main methods for predicting sea temperature. It is important to note that both methods have their advantages and disadvantages. Numerical ocean modeling considers comprehensive physical mechanisms and employs various methods to reduce input errors and improve forecast accuracy. On the other hand, data-driven methods rely on observed data to improve forecast accuracy without considering physical mechanisms.

Considering extensive physical mechanisms and using numerical algorithms to solve governing equations, ocean modeling is widely used to simulate or predict the ocean state [53]. These traditional numerical ocean models are usually computationally expensive and require various input data as complex geometric boundaries. Thus, accurate prediction of SST using a numerical ocean model remains a challenge due to the complex dynamical and thermal processes at the air–sea interface, such as ocean waves [54], turbulence, and radiation fluxes [55]. In general, accurate SST prediction by numerical ocean models usually requires input data, including topography, initial temperature and salinity fields, bottom and lateral boundary conditions, and atmospheric forcing fields, and it usually improves accuracy by data assimilation. For example, Lin used POM to predict the three-dimensional temperature field in the Bohai Sea, Yellow Sea, and East China Sea [56]. The Levitus climate monthly mean temperature and salinity were used as the initial fields, and model results from MOM2 were used as boundary conditions, respectively. The nudging assimilation method was also used to improve the accuracy of temperature prediction. As a result, the RMS between the temperature simulated by heat flux and the observed data are 1.56 °C, 0.88 °C, and 1.23 °C, respectively [57]. The atmospheric data and model parameter scheme also affected the simulation result even though the model have the same initial and boundary conditions. Gao et al. used HYCOM to simulate the sea surface temperature of the Tropical and North Pacific [13]. Based on the different combinations of two air–sea flux data sets (COADS and ECMWF) and two bulk parameter formulas (non-constant and constant), they found that the largest difference is about 4 °C using different combinations [13]. Compared with numerical ocean models, the topography, the boundary condition, and the physical parameter schemes can be ignored or disregarded by data-driven methods.

Many Different data-driven methods are used to predict the SST in the East China Sea [16,17,34,57–59]. However, this data-driven method only uses a single SST as input data and does not consider the influence of other factors affecting the sea surface temperature on the model prediction. The model proposed in this paper has preliminarily explored the influence of atmospheric factors on SST prediction. The SA-ConvLSTM used in this paper achieved good prediction results using the SST-only dataset, with the RMSE and MAE of about 0.72 °C and 0.56 °C, respectively. The atmospheric data for T2m and LWR are not used as model forcing conditions but as training and input data. Both the SST-T2m and SST-LWR datasets can improve model performance, but when the combined SST-T2m-LWR dataset is used, the prediction skill of the model improves significantly, with a 10.75% reduction in RMSE and a 9.06% reduction in MAE. The SA-ConvLSTM has the largest error in summer and the lowest error in other seasons. It indicated that the addition of T2m and LWR significantly improves the model's performance in summer while only slightly improving it in other seasons. The experiment results indicate that the accuracy of predicting the SST can be improved using more related data in the data-driven method.

This study shows that the Random Forest importance scores and R-value would help to select key influence factors to improve model performance. However, further exploration is needed to understand why the model has a larger MAE and RMSE in summer compared to the other three seasons. There are several possible solutions to this problem. Expanding the temporal span of data would be one solution way. The data used in this paper only covers the period from 2010 to 2020 and does not include signals of long-term trends, such as PDO cycles or global warming. It is necessary to experimentally test whether the model can be more accurate when using much longer data, such as 30 years or more.

Adjusting the parameters and settings of the model is another way to improve prediction accuracy. For example, modifying the length of input data can lead to more reasonable results. Additionally, integrating data-driven methods with the physical mechanisms that affect SST could be a potential solution, which requires further exploration.

6. Conclusions

This study aims to predict SST in the East China Sea using the Random Forest and SA-ConvLSTM model. The contribution of several key atmospheric forcings to the SST prediction has been tested. Based on the Random Forest importance assessment and the correlation coefficient R, the 2 m air temperature and longwave radiation were identified as the two most important additional data that can affect the prediction results of the SA-ConvLSTM model. Four datasets were constructed as input to SA-ConvLSTM: SST-only, SST-T2m, SST-LWR, and SST-T2m-LWR. Using all four datasets can represent the SST in the ECS well while having the highest RMSE and AE in summer and the lowest RMSE and AE in winter.

Compared with the prediction results using the SST-only dataset, when using the SST-T2m and SST-LWR, the prediction skill of the model can be improved by about 9.9% and 9.43% for the RMSE and by about 8.97% and 8.21% for the MAE, respectively. Using the SST-T2m-LWR dataset, the model's prediction skill can be improved by 10.75% for RMSE and 9.06% for MAE. Based on the Random Forest and the correlation coefficient R, the selection of the two main atmospheric forces can significantly improve the model's prediction performance. The model shows higher MAEs and RMSEs in summer compared to the other three seasons. More data with a longer time span, adjustment of model parameters and settings, and integration of data-driven methods with physical mechanisms could be used in further studies.

Author Contributions: Conceptualization, X.J., Y.W., X.L. and Q.J.; methodology, X.J. and Q.J.; software, X.J.; validation, X.J., Q.J. and L.J.; formal analysis, X.J., Q.J. and M.X.; investigation, X.J. and Q.J.; resources, Q.J. and X.L.; data curation, M.X., Z.M., X.L. and X.J; writing—original draft preparation, X.J.; writing—review and editing, Q.J. and X.J.; visualization, X.J.; supervision, Q.J.; project administration, Q.J. and L.J.; funding acquisition, Q.J. and L.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the construction Project of China ASEAN (Association of Southeast Asian Nations) blue partnership (No.99950410), National Natural Science Foundation of China under grant No. 41806004, and the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (SML2020SP007 and 311020004).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: OSTIA data sets can be found here: https://www.ncei.noaa.gov/data/oceans/ghrsst/L4/GLOB/UKMO/OSTIA/ (accessed on 11 November 2021). ERA5 data sets can be found here: https://cds.climate.copernicus.eu/ (accessed on 20 December 2021).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Herbert, T.D.; Peterson, L.C.; Lawrence, K.T.; Liu, Z. Tropical Ocean Temperatures over the Past 3.5 Million Years. *Science* **2010**, *328*, 1530–1534. [CrossRef] [PubMed]
2. Bouali, M.; Sato, O.T.; Polito, P.S. Temporal trends in sea surface temperature gradients in the South Atlantic Ocean. *Remote Sens. Environ.* **2017**, *194*, 100–114. [CrossRef]
3. Yao, S.-L.; Luo, J.-J.; Huang, G.; Wang, P. Distinct global warming rates tied to multiple ocean surface temperature changes. *Nat. Clim. Chang.* **2017**, *7*, 486–491. [CrossRef]
4. Lin, X.; Shi, J.; Jiang, G.; Liu, Z. The influence of ocean waves on sea surface current field and sea surface temperature under the typhoon background. *Mar. Sci. Bull.* **2018**, *37*, 396–403. [CrossRef]
5. Nagasoe, S.; Kim, D.-I.; Shimasaki, Y.; Oshima, Y.; Yamaguchi, M.; Honjo, T. Effects of temperature, salinity and irradiance on the growth of the red tide dinoflagellate Gyrodinium instriatum Freudenthal et Lee. *Harmful Algae* **2006**, *5*, 20–25. [CrossRef]

6. Li, S.; Goddard, L.; DeWitt, D.G. Predictive Skill of AGCM Seasonal Climate Forecasts Subject to Different SST Prediction Methodologies. *J. Clim.* **2008**, *21*, 2169–2186. [CrossRef]
7. Solanki, H.U.; Bhatpuria, D.; Chauhan, P. Signature analysis of satellite derived SSHa, SST and chlorophyll concentration and their linkage with marine fishery resources. *J. Mar. Syst.* **2015**, *150*, 12–21. [CrossRef]
8. Jiao, N.; Zhang, Y.; Zeng, Y.; Gardner, W.D.; Mishonov, A.V.; Richardson, M.J.; Hong, N.; Pan, D.; Yan, X.-H.; Jo, Y.-H.; et al. Ecological anomalies in the East China Sea: Impacts of the Three Gorges Dam? *Water Res.* **2007**, *41*, 1287–1293. [CrossRef] [PubMed]
9. Patil, K.; Deo, M.C. Prediction of daily sea surface temperature using efficient neural networks. *Ocean Dyn.* **2017**, *67*, 357–368. [CrossRef]
10. Wei, X.; Xiang, Y.; Wu, H.; Zhou, S.; Sun, Y.; Ma, M.; Huang, X. AI-GOMS: Large AI-Driven Global Ocean Modeling System. *arXiv* **2023**, arXiv:2308.03152. [CrossRef]
11. Stockdale, T.N.; Balmaseda, M.A.; Vidard, A. Tropical Atlantic SST Prediction with Coupled Ocean–Atmosphere GCMs. *J. Clim.* **2006**, *19*, 6047–6061. [CrossRef]
12. Krishnamurti, T.; Chakraborty, A.; Krishnamurti, R.; Dewar, W.; Clayson, C. Seasonal Prediction of Sea Surface Temperature Anomalies Using a Suite of 13 Coupled Atmosphere–Ocean Models. *J. Clim.* **2006**, *19*, 6069–6088. [CrossRef]
13. Song, G.; Xianqing, L.; Wang, H. Sea Surface Temperature Simulation of Tropical and North Pacific Basins Using a Hybrid Coordinate Ocean Model (HYCOM). *Mar. Sci. Bull.* **2008**, *10*, 1–14.
14. Zhang, X.; Zhang, W.; Li, Y. Characteristics of the sea temperature in the North Yellow Sea. *Mar. Forecast.* **2015**, *32*, 89–97.
15. Qian, Z.; Tian, J.; Cao, C.; Wang, Q. The numerical simulation and the assimilation technique of the current and the temperature field in the Bohai Sea, the Huanghai Sea and the East China Sea. *Haiyang Xuebao* **2005**, *27*, 1–6.
16. Xiao, C.; Chen, N.; Hu, C.; Wang, K.; Xu, Z.; Cai, Y.; Xu, L.; Chen, Z.; Gong, J. A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data. *Environ. Model. Softw.* **2019**, *120*, 104502. [CrossRef]
17. Wie, L.; Guan, L.; Qu, L.; Li, L. Prediction of Sea Surface Temperature in the South China Sea by Artificial Neural Networks. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 8158–8161. [CrossRef]
18. Garcia-Gorriz, E.; Garcia-Sanchez, J. Prediction of sea surface temperatures in the western Mediterranean Sea by neural networks using satellite observations. *Geophys. Res. Lett.* **2007**, *34*, L11603. [CrossRef]
19. Patil, K.; Deo, M.C. Basin-Scale Prediction of Sea Surface Temperature with Artificial Neural Networks. *J. Atmos. Ocean. Technol.* **2018**, *35*, 1441–1455. [CrossRef]
20. Zheng, G.; Li, X.; Zhang, R.H.; Liu, B. Purely satellite data-driven deep learning forecast of complicated tropical instability waves. *Sci. Adv.* **2020**, *6*, eaba1482. [CrossRef]
21. Xue, Y.; Leetmaa, A. Forecasts of tropical Pacific SST and sea level using a Markov model. *Geophys. Res. Lett.* **2000**, *27*, 2701–2704. [CrossRef]
22. Laepple, T.; Jewson, S.; Meagher, J.; O'Shay, A.; Penzer, J. Five year prediction of Sea Surface Temperature in the Tropical Atlantic: A comparison of simple statistical methods. *arXiv* **2007**, arXiv:physics/0701162. [CrossRef]
23. Kug, J.-S.; Kang, I.-S.; Lee, J.-Y.; Jhun, J.-G. A statistical approach to Indian Ocean sea surface temperature prediction using a dynamical ENSO prediction. *Geophys. Res. Lett.* **2004**, *31*, L09212. [CrossRef]
24. Peng, Y.; Wang, Q.; Yuan, C.; Lin, K. Review of Research on Data Mining in Application of Meteorological Forecasting. *J. Arid Meteorol.* **2015**, *33*, 19–27. Available online: http://www.ghqx.org.cn/EN/10.11755/j.issn.1006-7639(2015)-01-0019 (accessed on 13 October 2022).
25. Chaudhari, S.; Balasubramanian, R.; Gangopadhyay, A. Upwelling Detection in AVHRR Sea Surface Temperature (SST) Images using Neural-Network Framework. In Proceedings of the IGARSS 2008—2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; pp. IV-926–IV-929. [CrossRef]
26. Wu, Z.; Jiang, C.; Conde, M.; Deng, B.; Chen, J. Hybrid improved empirical mode decomposition and BP neural network model for the prediction of sea surface temperature. *Ocean Sci.* **2019**, *15*, 349–360. [CrossRef]
27. Tangang, F.T.; Hsieh, W.W.; Tang, B. Forecasting regional sea surface temperatures in the tropical Pacific by neural network models, with wind stress and sea level pressure as predictors. *J. Geophys. Res. Oceans* **1998**, *103*, 7511–7522. [CrossRef]
28. Wu, A.; Hsieh, W.W.; Tang, B. Neural network forecasts of the tropical Pacific sea surface temperatures. *Neural Netw.* **2006**, *19*, 145–154. [CrossRef] [PubMed]
29. Tripathi, K.C.; Das, M.L.; Sahai, A. Predictability of sea surface temperature anomalies in the Indian Ocean using artificial neural networks. *Indian J. Mar. Sci.* **2006**, *35*, 210–220.
30. Aparna, S.G.; D'Souza, S.; Arjun, N.B. Prediction of daily sea surface temperature using artificial neural networks. *Int. J. Remote Sens.* **2018**, *39*, 4214–4231. [CrossRef]
31. Gupta, S.M.; Malmgren, B. Comparison of the accuracy of SST estimates by artificial neural networks (ANN) and other quantitative methods using radiolarian data from the Antarctic and Pacific Oceans. *Earth Sci. India* **2009**, *2*, 52–75.
32. Hou, S.; Li, W.; Liu, T.; Zhou, S.; Guan, J.; Qin, R.; Wang, Z. MIMO: A Unified Spatio-Temporal Model for Multi-Scale Sea Surface Temperature Prediction. *Remote Sens.* **2022**, *14*, 2371. [CrossRef]
33. Zhang, Q.; Wang, H.; Dong, J.; Zhong, G.; Sun, X. Prediction of Sea Surface Temperature Using Long Short-Term Memory. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1745–1749. [CrossRef]

34. Xiao, C.; Chen, N.; Hu, C.; Wang, K.; Gong, J.; Chen, Z. Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sens. Environ.* **2019**, *233*, 111358. [CrossRef]
35. He, Q.; Hu, Z.; Xu, H.; Song, W.; Du, Y. Sea Surface Temperature Prediction Method Based on Empirical Mode Decomposition-Gated Recurrent Unit Model. *Laser Optoelectron. Prog.* **2021**, *58*, 9. [CrossRef]
36. Ham, Y.-G.; Kim, J.-H.; Luo, J.-J. Deep learning for multi-year ENSO forecasts. *Nature* **2019**, *573*, 568–572. [CrossRef] [PubMed]
37. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y.; Wong, W.-K.; Woo, W.-C. Convolutional LSTM Network: A machine learning approach for precipitation nowcasting. In Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Volume 1, pp. 802–810. [CrossRef]
38. Zhang, K.; Geng, X.; Yan, X.H. Prediction of 3-D Ocean Temperature by Multilayer Convolutional LSTM. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1303–1307. [CrossRef]
39. Luo, W.; Li, Y.; Urtasun, R.; Zemel, R. Understanding the effective receptive field in deep convolutional neural networks. In Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 4905–4913. [CrossRef]
40. Lin, Z.; Li, M.; Zheng, Z.; Cheng, Y.; Yuan, C. Self-Attention ConvLSTM for Spatiotemporal Prediction. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 11531–11538. [CrossRef]
41. Donlon, C.J.; Martin, M.; Stark, J.; Roberts-Jones, J.; Fiedler, E.; Wimmer, W. The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.* **2012**, *116*, 140–158. [CrossRef]
42. Stewart, R.H. *Introduction to Physical Oceanography*; Prentice Hall: Upper Saddle River, NJ, USA, 2008; pp. 51–52. [CrossRef]
43. Breiman, L. Bagging Predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
44. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
45. Belgiu, M.; Drăguţ, L.D. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
46. Chen, S.; Yu, L.; Zhang, C.; Wu, Y.; Li, T. Environmental impact assessment of multi-source solid waste based on a life cycle assessment, principal component analysis, and random forest algorithm. *J. Environ. Manag.* **2023**, *339*, 117942. [CrossRef] [PubMed]
47. Morse-McNabb, E.M.; Hasan, M.F.; Karunaratne, S. A Multi-Variable Sentinel-2 Random Forest Machine Learning Model Approach to Predicting Perennial Ryegrass Biomass in Commercial Dairy Farms in Southeast Australia. *Remote Sens.* **2023**, *15*, 2915. [CrossRef]
48. Wang, Y.; Chen, X.; Wang, Z.; Gao, M.; Wang, L. Integrating SBAS-InSAR and Random Forest for Identifying and Controlling Land Subsidence and Uplift in a Multi-Layered Porous System of North China Plain. *Remote Sens.* **2024**, *16*, 830. [CrossRef]
49. Su, H.; Li, W.; Yan, X.-H. Retrieving temperature anomaly in the global subsurface and deeper ocean from satellite observations. *J. Geophys. Res. Oceans* **2018**, *123*, 399–410. [CrossRef]
50. Cui, H.; Tang, D.; Mei, W.; Liu, H.; Sui, Y.; Gu, X. Predicting tropical cyclone-induced sea surface temperature responses using machine learning. *Geophys. Res. Lett.* **2023**, *50*, e2023GL104171. [CrossRef]
51. Chen, C.-T.A.; Sheu, D.D. Does the Taiwan Warm Current originate in the Taiwan Strait in wintertime? *J. Geophys. Res. Oceans* **2006**, *111*, C04005. [CrossRef]
52. Chen, C.-T.A. Rare northward flow in the Taiwan Strait in winter: A note. *Cont. Shelf Res.* **2003**, *23*, 387–391. [CrossRef]
53. Sonnewald, M.; Lguensat, R.; Jones, D.C.; Dueben, P.D.; Brajard, J.; Balaji, V. Bridging observations, theory and numerical simulation of the ocean using machine learning. *Environ. Res. Lett.* **2021**, *16*, 073008. [CrossRef]
54. Qiao, F.; Yuan, Y.; Yang, Y.; Zheng, Q.; Xia, C.; Ma, J. Wave-induced mixing in the upper ocean: Distribution and application to a global ocean circulation model. *Geophys. Res. Lett.* **2004**, *31*, L11303. [CrossRef]
55. Huang, C.J.; Qiao, F.; Chen, S.; Xue, Y.; Guo, J. Observation and Parameterization of Broadband Sea Surface Albedo. *J. Geophys. Res. Oceans* **2019**, *124*, 4480–4491. [CrossRef]
56. Lin, J. Numerical Simulation of Three-Dimensional Current Field and Temperature Field of the Bohai Sea, Yellow Sea and the East China Sea. Master's Thesis, Ocean University of China, Qingdao, China, 2004. Available online: https://www.dissertationtopic.net/doc/1278630 (accessed on 18 January 2024).
57. Jia, X.; Ji, Q.; Han, L.; Liu, Y.; Han, G.; Lin, X. Prediction of Sea Surface Temperature in the East China Sea Based on LSTM Neural Network. *Remote Sens.* **2022**, *14*, 3300. [CrossRef]
58. Xu, S.; Dai, D.; Cui, X.; Yin, X.; Jiang, S.; Pan, H.; Wang, G. A deep leaning approach to predict sea surface temperature based on multiple modes. *Ocean Modell.* **2023**, *181*, 102158. [CrossRef]
59. Yu, X.; Shi, S.; Xu, L.; Liu, Y.; Miao, Q.; Sun, M. A Novel Method for Sea Surface Temperature Prediction Based on Deep Learning. *Math. Probl. Eng.* **2020**, *2020*, 6387173. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.